

Diplôme Inter-Universitaire de Pédagogie Médicale

Université Pierre et Marie Curie – Paris 6

Application du Test de Concordance de Script à une évaluation de la compétence clinique en Orthopédie Pédiatrique
--

Raphaël VIALLE

Année Universitaire 2008-2009

Mémoire soutenu le 15 octobre 2009 à Paris

PLAN

Première partie : L'enseignement national du DESC de Chirurgie Pédiatrique

La formation théorique

La formation pratique

Evaluation et validation

Seconde partie : Le Test de Concordance de Script

Problématique

Caractéristiques du TCS

L'énoncé

Le recueil de la réponse

Modalités de scorage

Annexe 1: Canevas de questions et de grilles de correction

Troisième partie : Application du Test de Concordance de Script à une évaluation de la compétence clinique en Orthopédie Pédiatrique

1. Objectifs de l'étude

2. Méthodes

Définition des objectifs d'évaluation

Identification des situations cliniques pertinentes

Annexe 2 : Table de Spécification du TCS

Population-test

Validation du Panel de Référence

Analyse

3. Résultats

4. Discussion

5. Références

Résumé

Le Collège Hospitalo-Universitaire de Chirurgie Pédiatrique (CHUCP) a mené une réorganisation de l'enseignement national du DESC, conduisant à envisager la réalisation d'un examen national validant le DESC dont la première session aura lieu en 2009. Le Collège a décidé le recours aux Test de concordance de script (TCS) en complément d'autres modalités d'évaluation pour la partie écrite et théorique de la validation.

L'objectif de ce travail est la réalisation et la validation d'un test explorant les connaissances en orthopédie pédiatrique. Nous avons construit en concertation avec les orthopédistes pédiatres membres du collège un TCS d'orthopédie pédiatrique après identification de situations clinique courantes et problématiques. Ce test a été soumis à une population de chirurgiens pédiatres confirmés ou en cours de formation. L'analyse des scores montre que le TCS élaboré est très efficace pour discriminer des sous-groupes de compétences et d'expériences différents. Le TCS nous semble donc être un outil utilisable pour explorer l'organisation des connaissances en orthopédie pédiatrique. Le test élaboré va être ajouté à la banque nationale de questions du Collège Hospitalo-Universitaire de Chirurgie Pédiatrique.

Avant-propos

Ce travail reprend la structure et la forme d'un précédent travail réalisé par Marc-David LECLAIR :

Application du Test de Concordance de Script à une évaluation de la compétence clinique en Chirurgie Pédiatrique Viscérale. Projet d'évaluation de la progression du raisonnement clinique en Urologie Pédiatrique. Mémoire de DIU de pédagogie Médicale, Brest, 2007.

Le test de concordance de script d'orthopédie pédiatrique, la méthodologie de réalisation de l'étude et les résultats obtenus sont en revanche des éléments originaux du travail présenté ici.

Première partie : L'enseignement national du DESC de Chirurgie Pédiatrique

Le Collège hospitalo-universitaire de chirurgie pédiatrique a mené une réorganisation de l'enseignement du DESC de chirurgie infantile. L'objectif est d'assurer une formation de qualité aux chirurgiens pédiatres, situant l'enseignement de la chirurgie pédiatrique à un niveau comparable à ce qui est réalisé en Europe dans la spécialité et en France dans d'autres spécialités chirurgicales. L'organisation de cette formation tient compte du nombre de chirurgiens pédiatres en formation et de leur activité future partiellement spécialisée (chirurgie orthopédique ou viscérale pédiatrique). Cette formation de troisième cycle comprend la formation théorique, l'apprentissage au raisonnement clinique et la formation pratique.

La formation théorique :

Alors que les textes réglementaires prévoient un enseignement théorique organisé au niveau inter-régional, il est apparu souhaitable que cet enseignement soit réalisé à l'échelle nationale, comme dans d'autres spécialités chirurgicales (chirurgie plastique, chirurgie orthopédique et traumatologique), pour des raisons d'efficacité et d'homogénéité. Cet enseignement est organisé sous forme de modules, certains étant communs à tous les chirurgiens pédiatres en formation, d'autres spécifiques à l'orthopédie ou à la chirurgie viscérale pédiatrique. Les modules communs correspondent au socle de connaissance que doit avoir tout futur chirurgien pédiatre, quelle que soit son orientation. Ces modules communs représentent 36 heures d'enseignement répartis sur les trois années de formation. La majorité des modules sont donc des modules spécifiques. Concernant la chirurgie viscérale, ces modules représentent 84 h, répartis sur vingt-deux demi-journées. Ces modules sont répartis sous la forme de 2 sessions annuelles de deux et trois jours, deux fois par an. Un à trois enseignants sont responsables de la coordination d'un module. Ils s'adjoindront pour la réalisation de ce module les autres enseignants de leur choix, qu'ils appartiennent à la spécialité ou à d'autres spécialités. L'enseignement est dispensé sous la forme de courts magistraux. A l'issue de chaque module, l'étudiant reçoit les cours dispensés sur un CD-ROM. Les cours sont parallèlement mis en ligne sur le site Internet de la Société Française de Chirurgie Pédiatrique (SFCP) et de la Société Française d'Orthopédie Pédiatrique (SOFOP).

Cet enseignement national est destiné aux étudiants inscrits aux DESC de chirurgie infantile. La validation de l'ensemble des modules est obligatoire et devra être réalisée au cours des trois dernières années d'internat et des deux premières années de post internat. Ce caractère obligatoire s'impose, à partir de la rentrée universitaire 2005 – 2006, pour les internes commençant leur cinquième semestre. Les étudiants qui le souhaitent peuvent valider des modules durant leur deux premières années d'internat. Les étudiants plus avancés dans leur cycle de formation peuvent

assister à l'enseignement théorique mais il n'aura alors pas de caractère obligatoire.

L'apprentissage du raisonnement clinique :

L'apprentissage du raisonnement clinique est organisé dans chacune des inter-régions par le coordonnateur interrégional, assisté d'un vice coordonnateur de l'autre partie de la spécialité. Cet apprentissage se fait à l'occasion de staffs interrégionaux, organisés au minimum deux fois par an. Ces staffs sont des réunions communes ou distinctes de chirurgie viscérale et de chirurgie orthopédique pédiatrique. Chaque interne DESC présente au moins un dossier. Le dossier est complété par une analyse brève de la littérature sur les problématiques posées. Les enseignants de l'inter-région commenteront la présentation du dossier. Ces réunions interrégionales préparent l'étudiant à l'épreuve qui lui sera demandée lors de l'examen inter-régional.

La formation pratique :

Le suivi de la formation pratique dans chaque inter-région est assuré par le coordonnateur interrégional. A l'issue de chaque semestre, celui-ci reçoit, comme les textes le prévoient, deux fiches d'évaluation. La première est la fiche d'évaluation de l'interne par le chef de service, l'autre est la fiche d'évaluation du stage par l'étudiant. Cette fiche d'évaluation par l'étudiant est rajoutée au livret de l'interne et correspondra au document figurant en annexe de la circulaire du 2 septembre 2004. Ces 2 fiches seront élaborées à la fin du semestre en commun par l'interne et le chef de service, lors d'une réunion d'évaluation. Ces documents permettent au coordonnateur interrégional de suivre la formation de chaque étudiant et de dépister les difficultés, qu'il s'agisse d'une formation pratique insuffisante dans certains services ou d'étudiants qui pourraient paraître peu adaptés à l'exercice ultérieur de la spécialité.

Evaluation et validation :

Une épreuve de dossier validant l'apprentissage du raisonnement clinique est organisée à la fin de l'internat au sein de chaque inter-région comme stipulé par les textes réglementaires. L'évaluation de la formation théorique sera assurée par un examen national après un an de post-internat. Les membres du jury de l'examen national annuel seront les membres du conseil d'administration du collège, c'est-à-dire les coordonnateurs interrégionaux et les 2 membres du CNU. Pour l'examen, les coordonnateurs inter-régionaux seront assistés par les vice coordonnateurs de l'autre partie de la spécialité. Le jury pourra se répartir en deux ou trois sous-groupes. Dans ce cas, pour chaque étudiant, le jury comportera un membre de son inter-région et 2 membres d'une inter-région différente. Les étudiants pourront s'inscrire à l'examen du DESC, à l'issue de leur première année de post internat, lorsque leur maquette aura été validée et qu'ils auront suivi la totalité des modules.

L'inscription à l'examen national, alors que un module n'aurait pas été validé, ne serait possible que dans des circonstances exceptionnelles, jugée comme telles par le conseil d'administration du collège. L'inscription à l'examen comportera comme condition obligatoire d'avoir été dans les 3 premiers auteurs d'un article accepté dans une revue indexée. L'évaluation se rapprochera de l'épreuve du board européen et comportera une épreuve écrite, et la présentation du curriculum vitae et d'un mémoire (distinct de la thèse d'exercice). Ces règles d'examen du DESC s'appliqueront à partir de 2009, pour les DESC ayant pris leur première inscription avant le 1^{er} novembre 2005.

Le Collège hospitalo-universitaire de chirurgie pédiatrique a décidé de recourir au test de concordance de script pour l'épreuve écrite de l'examen national annuel. A l'origine de cette décision, un travail important a été réalisé par l'équipe de chirurgie pédiatrique du CHU de Rouen (Pr J Le Chevallier, Pr A Liard, Pr L Sibert), en particulier en ce qui concerne la partie orthopédique de la validation. Concernant la partie urologique, une validation de la pertinence de ces tests a été obtenue à l'issue du travail du Dr M Leclair (Brest) rentrant dans le cadre de son DIU de Pédagogie médicale.

L'objectif du présent travail est de contribuer à l'élaboration et à la mise en place de l'épreuve écrite de l'examen national annuel validant l'enseignement théorique du DESC de chirurgie pédiatrique, en élaborant et en validant un test explorant les connaissances en orthopédie pédiatrique.

Seconde partie : Le Test de Concordance de Script

Problématique

Les professionnels de santé rencontrent dans leur pratique quotidienne des problèmes de complexité variable. Pour les résoudre, ils font appel à leurs connaissances, mais aussi à leur expérience et à leur raisonnement clinique. Ces problèmes sont de deux sortes. Certains sont des problèmes bien définis, avec des données précises, stables, des buts clairs et prévisibles. Leur résolution exige la capacité d'appliquer des connaissances, la compétence est d'ordre technique. En pratique clinique, les médecins sont pourtant souvent confrontés à ce que l'on nomme des problèmes mal définis : les informations disponibles sont incomplètes, parfois ambiguës, le but à atteindre n'est pas toujours clair, les données sont incertaines ou peuvent faire l'objet de différentes interprétations. Au-delà de la maîtrise des différents savoirs techniques, cette capacité à raisonner en contexte d'incertitude est au centre de la compétence professionnelle.

Les outils de mesure par écrit du raisonnement clinique sont souvent imparfaits, en particulier parce

qu'ils ne permettent pas une bonne discrimination en fonction de la compétence. Une des raisons peut être que seuls des problèmes bien définis sont généralement soumis, alors que les compétences professionnelles sont principalement utilisées dans la résolution de problèmes mal définis. Les tâches de résolution de problèmes mal définis entraînent une difficulté supplémentaire en matière de correction. Soumis à un problème mal défini, les réponses fournies par des experts peuvent diverger, reflétant ainsi la complexité et les incertitudes entourant la question. Il n'existe donc souvent pas une seule réponse univoque mais un faisceau de possibilités retenues par un panel d'experts. Le test de concordance de script (TCS) apporte une réponse à ces problématiques. Outil de mesure de la compétence clinique, le TCS est conçu pour évaluer la capacité de raisonner en contexte d'incertitude et de résoudre des problèmes mal définis. Il est novateur à trois niveaux : le stimulus soumis à l'étudiant (la vignette décrivant la situation clinique), le moyen d'enregistrement des réponses, et la méthode d'établissement des scores pour chaque question. Cet outil ne remplace pas mais vient en complément d'autres méthodes d'évaluation comme les QCM, qui explorent l'acquisition de compétences techniques.

Bases théoriques

Le développement de l'expertise chez les professionnels est très lié à l'organisation des connaissances. Le raisonnement clinique est un processus hypothético-déductif (Elstein and al. 1978) caractérisé par la génération précoce d'hypothèses (orientées par la collection d'information) et la prise de décision (par l'utilisation d'information permettant de confirmer ou de rejeter les hypothèses émises). De nombreux auteurs ont émis l'hypothèse qu'en médecine clinique, un praticien expérimenté diffère d'un collègue moins expérimenté par la mise en œuvre de réseaux élaborés de connaissance adaptés aux tâches qu'il effectue de façon régulière. Ces réseaux de connaissances pré-établis appelés scripts (Feltovich and al. 1984, Schmidt and al. 1990, Charlin and al. 2000) sont organisés pour atteindre des objectifs au sein des tâches concernées : diagnostic, stratégie d'investigation ou thérapeutique. Ces scripts se constituent dès que les étudiants sont confrontés à des tâches cliniques réelles et sont ensuite développés et enrichis et affinés au cours de la vie professionnelle. Ils sont utilisés pour comprendre et pour traiter activement l'information clinique. Cette théorie met en évidence que dans une situation clinique donnée, le clinicien fait appel à la connaissance mémorisée en rapport avec une situation clinique de même nature antérieurement vécue. La réactivation des connaissances à partir de la mémoire est ensuite utilisée dans un processus déductif pour activer la recherche d'information qui permettra la confirmation ou au contraire l'infirmerie des hypothèses évoquées (Charlin and al. 1994). L'élaboration de la connaissance (qui relève de la constitution de nombreux liens entre les différentes composantes des connaissances acquises et des expériences vécues) est la clef de la compétence clinique (Bordage

and al. 1994). Charlin and al. (1998) a conçu un instrument, le test de concordance de script, qui mesure le degré d'élaboration d'un schéma de diagnostic. Les principes du test de concordance de script (TCS) ont été déjà rapportés (Charlin and al. 2000). Cet instrument a été testé en radiologie, en gynécologie obstétrique et en urologie. Ces études ont mis en évidence que les seniors, les résidents, les étudiants ont des capacités de réponse différentes avec des niveaux de réponse différents avec l'expérience et des différences significatives entre chaque groupe (Charlin and al. 1998, Leclair, 2007). Ces différentes publications rappellent que le TCS a le double avantage d'être relativement simple à construire et d'être un outil de mesure fiable. Ces résultats apportent aussi de nouvelles modalités d'évaluation de la connaissance clinique. Beaucoup de compétences diagnostiques et thérapeutiques sont sous-évaluées par les enseignants. En effet, il est habituel d'évaluer les résidents et les chefs de clinique en appréciant de façon subjective leur performance par un examen oral sur un cas clinique isolé. La plupart de ces méthodes d'évaluation mesure une compétence clinique factuelle plus que l'organisation des connaissances. De nombreuses études ont montré que les évaluations objectives et structurées par des patients standardisés permettaient de détecter des différences dans la performance clinique entre des résidents en urologie sans différence en fonction de leur année d'étude (Sibert and al. 1997, 2000). Les limites de ce système d'évaluation sont bien connues en particulier les difficultés logistiques (préparation des candidats, préparation de l'équipe enseignante, locaux, préparation des patients simulés...) (Cohen and al. 1990, Grand'maison and al 1992).

Caractéristiques du TCS

L'énoncé

Egalement appelé vignette, l'énoncé décrit une situation clinique réaliste qui doit être problématique même pour un clinicien averti. Plusieurs options (diagnostique, prise en charge) sont envisageables, et la description peut être incomplète, et certaines données sont ambiguës ou conflictuelles. La tâche consiste à envisager l'effet que produirait l'apparition d'une nouvelle donnée sur le statut de l'une des options pertinentes, puis d'une autre donnée sur une autre option, etc. Pour une situation clinique donnée, plusieurs questions successives (les items) peuvent être envisagées, indépendantes les unes des autres. La structure du TCS est donc très stéréotypée, associant la description de la situation clinique, suivie par une série de questions qui comportent trois parties :

- une hypothèse diagnostique, une investigation complémentaire, ou une option thérapeutique envisageables, plausibles au vu de la situation clinique
- une nouvelle information

-une échelle de Likert en 5 points qui mesure l'effet de la nouvelle donnée sur la pertinence de l'hypothèse.

Le recueil de la réponse

Une originalité du TCS est l'utilisation d'une échelle graduée 3 ou 5 valeurs (échelle de Likert), dont la signification est précisée dans l'énoncé. Il s'agit de brèves propositions d'ordre qualitatif. Les différents points de l'échelle correspondent à des valeurs positives (l'option est renforcée par la nouvelle donnée), neutre (la donnée ne modifie pas le statut de l'option) ou négative (l'option est infirmée par la donnée).

Modalités de scorage

Le scorage est l'un des autres éléments très originaux du TCS. On considère que la réponse donnée à chaque question pour un clinicien donné n'est jamais de type binaire de sorte que, pour une question donnée, deux cliniciens expérimentés n'apporteront pas exactement la même réponse. Il est donc plus pertinent de réunir un panel d'experts, qui devra répondre individuellement aux questions du test dans les mêmes conditions que les étudiants testés. Pour chaque item, chaque réponse se voit affecter une valeur correspondant à la proportion d'experts qui ont choisi la réponse donnée. Par exemple, à une question donnée

10 % des experts ont répondu A,

30 % des experts ont répondu B,

60 % des experts ont répondu C,

aucun n'a répondu D ou E.

Le crédit sera alors attribué au prorata des réponses données en phase avec le groupe correspondant du panel:

l'étudiant qui répond A obtiendra 0,1 point (il a répondu comme 10% du panel),

l'étudiant ayant répondu B obtiendra 0,3 points (il a répondu comme 30% du panel)

l'étudiant ayant répondu C aura 0,6 points (il a répondu comme 60% du panel)

l'étudiant ayant répondu D ou E aura 0 point.

Enfin, le crédit accordé pour chaque réponse est pondéré de sorte que la note maximale soit 1 et la note minimale soit 0. Dans l'exemple choisi,

la réponse C rapportera 1 point (0,6 / 0,6),

la réponse B rapportera 0,5 points (0,3 / 0,6),

la réponse A rapportera 0,16 points (0,1 / 0,6),

les réponses C et D rapporteront 0 point.

Annexe 1: Canevas de questions et de grilles de correction

Le format dépend du type de compétence évaluée (Faire un diagnostic, Demander une exploration complémentaire, Choisir un traitement).

Exemple 1 : Evaluation d'une aptitude à proposer un diagnostic

Si vous pensez à...	Et que vous observez...	Votre hypothèse devient...				
<i>Une hypothèse diagnostique</i>	<i>Une information clinique, un cliché d'imagerie, un résultat d'examen complémentaire</i>	-2	-1	0	+1	+2

-2 Mon hypothèse est pratiquement éliminée

-1 Mon hypothèse devient peu probable

0 Cette information n'a aucun effet sur mon hypothèse diagnostique

+1 L'hypothèse devient plus probable

+2 Mon hypothèse devient pratiquement certaine

Exemple 2 : Evaluation d'une aptitude à choisir une investigation complémentaire

Si vous pensiez demander...	Et qu'alors vous trouvez	L'effet sur la nécessité de demander ce test est le suivant				
<i>Option d'examen para clinique</i>	<i>Information clinique ou paraclinique</i>	-2	-1	0	+1	+2

-2 absolument contre-indiqué

-1 peu utile ou plutôt néfaste

0 ni plus ni moins utile

+1 utile et souhaitable

+2 indispensable

Exemple 3 : Evaluation d'une compétence à proposer un choix thérapeutique

Si vous pensiez faire	Et qu'alors vous trouvez	Votre hypothèse devient...				
<i>Une option thérapeutique</i>	<i>Une nouvelle information clinique ou le résultat d'un examen complémentaire</i>	-2	-1	0	+1	+2

-2 absolument contre-indiqué

-1 peu utile ou plutôt néfaste

0 ni plus ni moins utile

+1 utile et souhaitable

+2 indispensable

Troisième partie : Application du Test de Concordance de Script à une évaluation de la compétence clinique en Orthopédie Pédiatrique

1. Objectifs de l'étude

Ce travail avait pour objectif d'établir la faisabilité d'un examen reposant sur l'utilisation des TCS pour l'évaluation de la compétence clinique en orthopédie pédiatrique. Le travail a comporté une phase de rédaction de l'examen, puis de validation. Aux cours de ces phases, nos objectifs étaient : - de déterminer les situations cliniques pertinentes pouvant faire l'objet d'une évaluation

-d'écrire une série complète de questions couvrant l'ensemble des situations cliniques identifiées, afin d'évaluer la faisabilité et les difficultés de rédaction

-d'évaluer la validité de construit de l'outil d'évaluation dans le champ clinique très spécifique de l'orthopédie pédiatrique, en soumettant les questions à un panel de participants d'expérience diverses

-d'améliorer par la discussion entre experts la qualité et la pertinence des items rédigés afin de les intégrer dans une banque nationale de questions.

2. Méthodes

Définition des objectifs d'évaluation

Comme explicité dans l'introduction, les candidats évalués sont des étudiants DESC de chirurgie infantile, à la fin de leur première année de post-internat. Il s'agit donc d'une population que l'on peut considérer comme homogène dans ses acquisitions et ses savoirs, et qui ne s'est pas encore particulièrement spécialisée dans son exercice.

L'évaluation à un objectif de certification, en complémentarité avec d'autres critères (présentation d'un mémoire original, épreuve orale de dossier clinique, présentation du CV). L'épreuve écrite combine les TCS avec des QCM à contexte riche évaluant plus spécifiquement les savoirs théoriques fondamentaux et les acquisitions de compétences techniques.

Identification des situations cliniques pertinentes

Au sein de notre thématique de travail, nous avons travaillé à identifier des situations cliniques pertinentes, soit par leur fréquence, soit par leur mode de présentation, leur gravité, ou leur possibilité de traitement. Nous avons ainsi isolé 38 situations cliniques problématiques (cf. Annexe 2), pour lesquelles nous avons retenus des domaines de connaissance spécifiques à

évaluer en priorité (diagnostic, examens complémentaires, traitement). Il s'agissait presque toujours de problèmes cliniques fondamentaux éducationnels envisagés dans le cadre du DESC, de type « motif de consultation », ou de situations cliniques caractéristiques des problèmes de raisonnement rencontrés dans la discipline. Ces situations cliniques et leurs domaines de compétences à évaluer sont résumées dans la table de spécification ci-après.

Annexe 2: Table de Spécification du TCS - ORTHOPEDIE PÉDIATRIQUE

SITUATION CLINIQUE	Nombre de TCS
Tumeurs malignes osseuses - traitement	2
Luxation congénitale de hanche	2
Tumeurs bénignes osseuses - diagnostic et traitement	3
Fracture supracondylienne	1
Fracture avant-bras et poignet	1
Fracture de jambe	1
Plaies de la main	1
Polytraumatisme de l'enfant	2
Croissance de l'appareil musculo-squelettique	2
Pathologie du genou en croissance	2
Handicap	2
Ostéomyélite	2
Pathologie du sport	2
Arthrite juvénile idiopathique	2
Ostéochondrite primitive de hanche	2
Scoliose neuro	1
ILMI	4
Boiteries	3
Spondylolyse et spondylolisthésis	2
Depistage LCH	1
Tumeurs bénignes osseuses - diagnostic et traitement	1
Ostéochondrite primitive de hanche	1
Cyphoses	2
Traumatismes craniens	2
Déformations des membres inférieurs dans le plan frontal	3
Epiphysiolyse	2
Paralysies du plexus brachial	2
Scolioses neurologiques	2
Tumeurs malignes osseuses - diagnostic	2
Scolioses idiopathiques	2
Pieds plats - pieds creux	3
Handicap	2
Arthrites bactériennes	3
Malformations des membres	1
Fragilités osseuses - diagnostic et traitement	2
Déformations des membres inférieurs dans le plan axial	3
Déformations des pieds d'origine congénitale (PBVE , PCC)	3
Déformations des pieds d'origine neurologique	3

Rédaction des vignettes et items

Les situations cliniques problématiques sont présentées sous la forme de petites vignettes, chacune suivie d'une série de questions (items). La vignette essaie de reproduire un contexte clinique authentique, dans un scénario qui pose problème même pour un expert. Les vignettes et items sont prévus pour être soumis sous forme de diapositives successives, incluant le cas échéant des documents iconographiques. Un examen comportant un total de 114 questions a été élaboré en réunion plénière des enseignants d'orthopédie pédiatrique du CHUCP. Une population-test a été soumise à cet examen en plusieurs séances. A chaque fois, une séance de « débriefing » a fait suite à l'étape de test proprement dit, afin de discuter la pertinence des items et retenir, de modifier ou d'éliminer les items qui intégreront la banque de questions.

Population-test

L'examen a été testé auprès d'une population de médecins de divers degrés d'expérience. Au total, 11 chirurgiens confirmés ou en cours de formation ont passé le test. Les participants ont été répartis en 3 groupes *a priori*, en fonction de leur cursus et de leur expérience clinique. Un groupe de niveau 1 (n=6), constitué d'internes DESC de chirurgie orthopédique en fin de cursus et un groupe de niveau 2 (n=2) composé de deux CCA en fin de cursus et un groupe de niveau 3 (n=3) (experts) constitué de praticiens hospitaliers d'Orthopédie Pédiatrique.

Validation du Panel de Référence

La validation du panel de référence (« experts ») s'est faite en 2 temps. Une première correction a été réalisée en prenant comme panel de référence trois praticiens hospitaliers expérimentés du groupe 2. Cette correction a permis d'établir un classement. Les autres participants recevant les meilleures notes ont été ajoutés au panel d'experts, permettant d'établir des simulations de corrections à partir d'un panel de référence comportant 3, 4, ou 5 « experts ». Il était intéressant de constater que les notes des participants variaient assez peu, et surtout que le classement des participants était identique en fonction des différentes simulations. Nous avons donc choisi de restreindre le panel de référence aux 3 experts désignés initialement.

Analyse

Les résultats de chaque groupe étaient exprimés sous la forme d'un score médian (extrêmes) et d'un score moyen ($\pm$ SD) sur 100. Les scores ont été calculés à l'aide d'une série de feuilles Microsoft Excel comportant des macros permettant une correction automatisée. Les scores moyens de chaque groupe ont été comparés par analyse de variance unifactorielle. Etant donnés les faibles effectifs dans chaque groupe et la possible non normalité des distributions, les résultats

ont également été comparés par un test de Kruskal-Wallis, qui est un équivalent non paramétrique d'une ANOVA à un facteur.

3. Résultats

Les résultats sont détaillés dans le Tableau 1. La médiane du score total était de 53,7 sur 100 (extrêmes 37.2 -93.3), et la moyenne de $58,4 \pm 11.4$. La durée requise pour la réalisation du test (incluant un préalable de 10 diapositives d'explications et d'exemples) et les 114 questions du test était d'environ 120 minutes. Le score moyen des participants de niveau 1 (internes DES, internes début DESC) était de $36,2 (\pm 4,3)$, celui des participants de niveau 2 (CCA) était de $72,3 (\pm 8,9)$, et celui des experts (niveau 3) était de $84,7 (\pm 11,0)$. L'analyse de variance montre que ces scores sont significativement différents entre eux ($dl=14$, $F=35,2$ $p < 10^{-3}$), ce qui est également confirmé par le Kruskal-Wallis ($p=10^{-2}$).

Ces résultats sont présentés graphiquement sur la figure 1 et en détail dans le tableau 1

Figure 1 : représentation graphique des résultats dans les trois groupes.

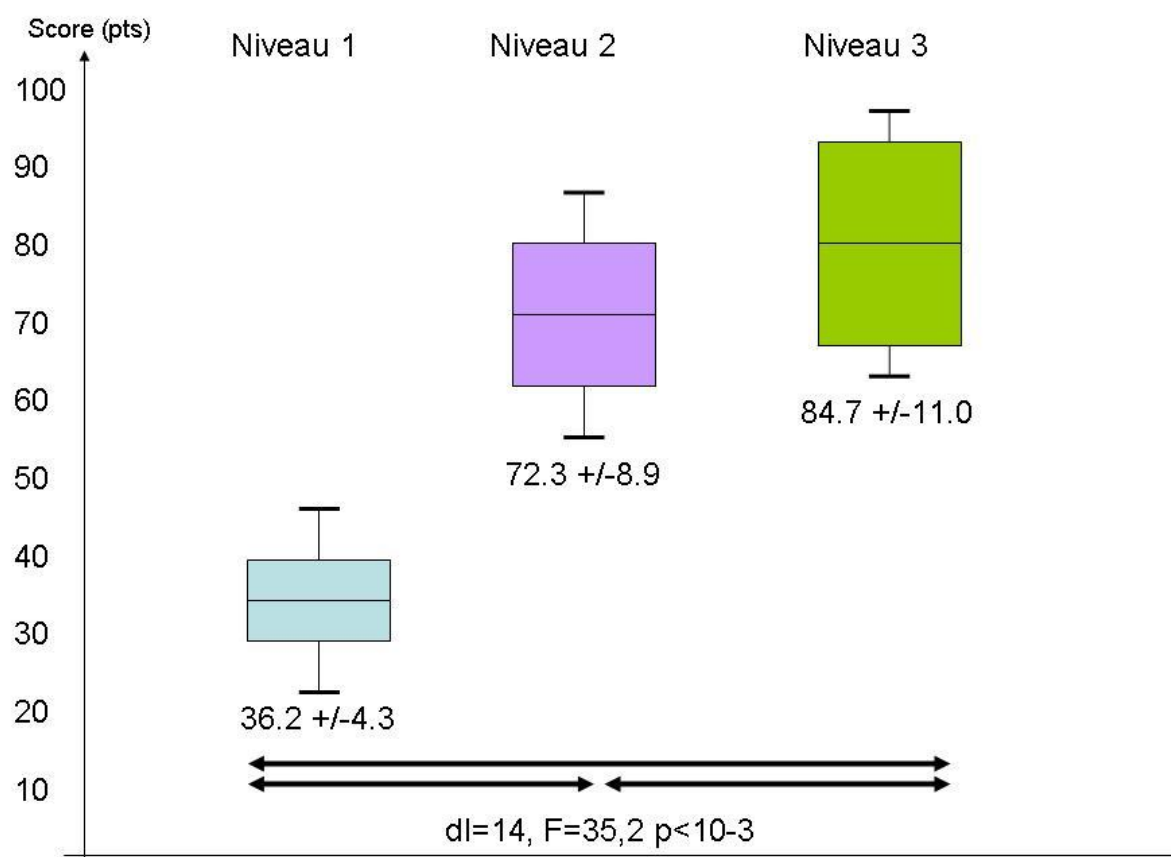


Tableau 1 : Résultats

médiane		moyenne	
n=		(extrêmes)	(± déviation standard)
Total	11	53,7 (29 -96)	58,4 (±11,4)
niveau 1	6	34,5 (29 – 56)	36,2 (±4,3)
niveau 2	2	67,3 (54 -73)	72,3 (±8,9)
niveau 3	3	81,7 (45 -96)	84,7 (±11,0)

4. Discussion

Ce travail préliminaire montre que les Tests de Concordance de Script sont utilisables en orthopédie pédiatrique.

Le principe du test de Concordance de Script s'appuie sur les théories du développement de l'expertise clinique qui soutiennent que l'expertise est en grande partie liée à l'organisation des connaissances et aux procédés cognitifs d'activation des connaissances pour faire face aux situations se présentant dans la pratique. Ce test a été développé dans le but d'explorer la capacité d'interpréter et de traiter au fur et à mesure les informations découvertes dans la résolution de problèmes cliniques. Ceci est clairement une composante de la compétence plus importante que la simple mémorisation de connaissances factuelles (Sibert, Prog Urol). Dans notre travail, les résultats obtenus par les participants augmentent avec l'expérience clinique, avec des scores clairement meilleurs dans le groupe de niveau 3 (praticiens hospitaliers) que dans le groupe de niveau 1 (internes DES). Ceci nous apparaît comme une validation *a posteriori* de la pertinence du test que nous avons élaboré, et confirme la validité de construit de l'outil d'évaluation, puisqu'il permet une bonne discrimination entre des groupes de sujets de niveaux de compétence clinique variés.

Les effectifs faibles dans chaque groupe représentent certainement une limite à l'interprétation de cette étude ; toutefois les différences observées étaient très nettes, et ne peuvent être expliquées par le seul hasard. Les résultats des groupes de niveau 1 et 2 sont probablement les plus intéressants à analyser, car ils représentent la cible de la population à évaluer lors de cet examen (DESC en fin de cursus et CCA). Nous avons observé dans ces groupes de bons résultats avec toutefois d'assez grandes disparités entre les candidats. Concernant le groupe des internes DESC, il est utile de préciser qu'il s'agit d'internes de Chirurgie Orthopédique et Traumatologique et non d'internes purement d'orthopédie pédiatrique. L'évaluation est toutefois intervenue à l'issue d'un stage de 6

mois de pratique exclusive de l'orthopédie pédiatrique.

Une difficulté lors de l'élaboration de ces tests est la constitution du panel de référence. Même si il est admis dans la définition-même du TCS que les réponses à un problème mal-défini ne sont pas toujours univoques, le nombre de bonnes réponses admissibles à un item donné nous paraît limité. Cela pose alors le problème de la définition et de la sélection de l'expert. Nous avons volontairement mis de coté ce problème en limitant le nombre d'experts dans notre panel de référence, considérant que notre objectif principal était de tester notre série de questions sur différents groupes participants pour vérifier le pouvoir discriminant du TCS. La constitution d'un panel de référence reste un écueil potentiellement important pour la généralisation de cet instrument à une échelle nationale et dans un but de certification.

Au-delà de la validation auprès d'un échantillon-test de cette technique d'évaluation, ce travail permet de débiter la constitution d'une banque de questions. Parmi les 114 questions rédigées, 12 semblent devoir être éliminées en raison de la grande dispersion des réponses proposées y compris par les candidats expérimentés.

Au total, le Test de Concordance de Script nous semble être un outil très intéressant dans l'évaluation de la compétence clinique en orthopédie pédiatrique. Le TCS nous paraît bien discriminer des groupes de participants de niveaux de compétence différents en orthopédie pédiatrique. Ce travail nous a permis d'élaborer, de tester, et de valider une série de questions d'orthopédie pédiatrique qui serviront à alimenter la banque de questions nationales du Collège Hospitalo-Universitaire de Chirurgie Pédiatrique

5. Références

BORDAGE, G. Elaborated knowledge: A key to successful diagnostic thinking. *Academic medicine*, 1994 69, pp. 883-885.

CHARLIN, B. Le schéma comme structure de connaissances sous-jacente aux hypothèses dans l'investigation clinique médicale. Mémoire pour l'obtention du diplôme de maîtrise ès arts en sciences de l'éducation, Université de Sherbrooke 1994.

CHARLIN, B., BRAILOVSKY, C.A., BRAZEAU-LAMONTAGNE, L., SAMSON, L., LEDUC, C. (1998a) Script questionnaires: their use for assessment of diagnostic knowledge in radiology *Medical Teacher*, 1998 20, pp.567-571.

CHARLIN, B., BRAILOVSKY, C. A., LEDUC, C., BLOUIN D. (1998b) The Diagnostic Script Questionnaire: A new tool to assess a specific dimension of clinical competence *Advances in Health Sciences Education*, 1998, 3, pp. 51-58.

CHARLIN, B., TARDIF, J., BOSHUIZEN, H. (2000a) Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. *Academic Medicine*, 2000, 75, pp.182-190.

CHARLIN, B., ROY, L., BRAILOVSKY, C. A., & VAN DER VLEUTEN, C. P. M. (2000). The Script Concordance Test: A tool to assess the reflective clinician. *Teaching and Learning in Medical Education*, 12, 189-195.

CHARLIN B, GAGNON R, SIBERT L, VAN DER VLEUTEN C (2002). Le test de concordance de script, un instrument d'évaluation du raisonnement clinique. *Pédagogie Médicale*, 3, pp 135-144

COHEN, R., RIEZNICK, R. K., TAYLOR, B. R., PROVAN, J., ROTHMAN, A. (1990) Reliability and validity of the objective structured clinical examination in assessing surgical residents *American journal of Surgery*, 160, pp. 302-305.

CUSTERS, J. F. M., REGHER, G., NORMAN, G. R. Mental representations of medical diagnostic knowledge: a review *Academic medicine*, 1996 71, pp. S55-S61.

ELSTEIN, A. S., SHULMAN, L. S., SPRAFKA, S. A. (1978) Medical problem solving: an analysis of clinical reasoning (Cambridge, MA, Harvard University Press)

FELTOVICH, P. J. (1983) Expertise: reorganizing and refining knowledge for use, Professions Education Research Notes, 4, pp. 5-9.

FELTOVICH, P. J., BARROWS, H. S. (1984) Issues of generality in medical problem solving. In H. G. Schmidt and M. L. De Volder (Eds.), Tutorials in problem-based learning: A new direction in teaching the health professions, Assen, Holland: Van Gorcum.

GRAND'MAISON, P., LESCOP, J., RAINSBERRY, P., BRAILOVSKY C. A. (1992) Large-scale use of an objective structured clinical examination for licensing family physicians Canadian Medical Association Journal, 146, pp. 1735-1740.

Mc NICOLL, A., BRAILOVSKY, C. A., BERTRAND, R., CARDINET, J. EtudGen, programme pour l'analyse de la généralisabilité pour Macintosh. CESSUL 1992, 1996, 1999. In D. BAIN et G. PINI. Pour évaluer vos évaluations: la généralisabilité, mode d'emploi. Centre de recherche psychopédagogiques, Genève, 1996.

LECLAIR MD, Application du Test de Concordance de Script à une évaluation de la compétence clinique en Chirurgie Pédiatrique Viscérale. Projet d'évaluation de la progression du raisonnement clinique en Urologie Pédiatrique. Mémoire de DIU de pédagogie Médicale, Brest, 2007

SCHMIDT, H. G., NORMAN, G. R., BOSHUIZEN, H. P. A. (1990) A cognitive perspective on medical expertise: Theory and implications Academic Medicine, 65, pp. 611-621.

SIBERT, L., GRAND'MAISON, P., CHARLIN, B., GRISE, P. (1997). Evaluation de la compétence clinique en urologie: approche innovatrice basée sur l'observation de la performance. Progrès en urologie, 7, pp. 581-589.

SIBERT, L., GRAND'MAISON, P., DOUCET, J., WEBER, J., & GRISE, P. (2000) Initial experience of an Objective Structured Clinical Examination in evaluating urology residents,

European Urology, 37, pp. 621-627.

SIBERT L., CHARLIN B., CORCOS J., GAGNON R., LECHEVALLIER J., P.GRISE P..

Assessment of clinical reasoning competence in urology with the Script Concordance test: An exploratory study across two sites from different countries. European Urology 2002;41:227-233.

SIBERT L., CHARLIN B., GAGNON R., CORCOS J., KHALAF A, P.GRISE P. (2001) Evaluation du raisonnement clinique en urologie : l'apport du test de concordance de script. Prog Urol 2001 ; 11 : 1213-1219

ZEITZ, C. M. (1997) Some concrete advantages of abstraction: how experts' representations facilitate reasoning. In P. J. FELTOVICH, K. KM. FORD, R. R. HOFFMAN (Eds.) Expertise in Context: Human and machine. Memlo Park, CA: AAAI / MIT press